세션 4

시각화 만들기

집계 개념 · 컴포넌트별 설정 · 국가별 접속 지도

01
집계 개념
Metric과 Bucket, Terms·Date Histogram

02
컴포넌트
Metric·Pie·Bar·Line·Table·Heat Map

03
Region Map
국가별 접속을 세계지도에 색으로 표현

04
저장·재사용
이니셜 규칙과 대시보드 재사용

S4
시각화

세션 개요와 학습 목표

> 시각화 컴포넌트와 집계를 익혀 보안 지표를 한눈에 드러낸다.

이 세션에서 얻는 것

- **컴포넌트 감각**: Metric·Pie·Bar·Line·Data Table·Heat Map·Region Map을 언제 어디에 쓰는지 구분한다.
- **집계 이해**: Terms와 Date Histogram으로 로그를 순위·추이로 요약하는 방법을 익힌다.
- **지표 시각화**: 공격 건수·유형·출발지·국가 분포를 그림으로 만들어 관제 화면 재료를 확보한다.

시각화의 목적

- **추이와 비율**: 시간대별 추이(Line)와 구성비(Pie)로 상황을 직관적으로 파악한다.
- **순위와 밀도**: 상위 공격 유형·IP 순위(Bar·Table)와 국가×상태코드 밀도(Heat Map)를 읽는다.

S4
시각화

Visualize 생성 흐름

> 모든 시각화는 같은 순서로 만들어지므로 흐름을 몸에 익힌다.

생성 단계

- **진입**: 좌측 메뉴에서 Visualize를 열고 Create visualization을 누른다.
- **선택**: 갤러리에서 컴포넌트를 고르고 인덱스 logs-demo-*를 대상으로 지정한다.
- **설정**: Metric(값)과 Bucket(축)을 지정한 뒤 상단 Update로 화면에 반영한다.
- **저장**: Save를 눌러 이름을 붙여 두면 나중에 대시보드에 다시 얹을 수 있다.

공용 환경 저장 규칙

- **이니셜 접두**: 공유 서버이므로 이름 앞에 본인 이니셜을 붙여 남의 시각화와 섞이지 않게 한다.
- **예시**: kim-공격유형처럼 이니셜-내용 형태로 저장해 목록에서 바로 찾는다.

S4
시각화

집계 개념 — Metric vs Bucket

> 시각화는 값을 계산하는 Metric과 나누는 Bucket의 조합이다.

두 축의 역할

- **Metric**: 무엇을 셀지 정하는 값 계산으로 Count·Sum·Avg·Max 같은 지표를 만든다.
- **Bucket**: 어떤 기준으로 나눌지 정하는 분류 축으로 X축·슬라이스·행이 된다.
- **조합 구조**: 보통 X축은 Bucket, Y축은 Metric으로 두어 분류별 값을 그린다.

대표 집계 종류

- **Terms**: 값별로 나눠 상위 N개 순위를 만든다(공격 유형·IP·국가).
- **Date Histogram**: 시간을 일정 간격으로 나눠 추이를 그린다(@timestamp).
- **Range·Filters**: 숫자 구간이나 조건별로 묶어 분류한다.

S4
시각화

Terms 집계와 .keyword 필수

> 텍스트 필드를 나눌 땐 반드시 .keyword를 선택해야 한다.

```
[올바른 Terms 설정]
Bucket:  X-axis
Aggregation:  Terms
Field:  attack_type.keyword
Order by:  Count
Order:  Descending
Size:  10
  → 상위 10개 공격 유형 순위

[잘못된 설정 – 오류 발생]
Field:  attack_type   (.keyword 없음)
  → "invalid for Terms aggregation"
  → text 필드는 그대로 집계 불가

[규칙]
텍스트 = 필드.keyword 로 집계
숫자·날짜 = 필드명 그대로
```

필드별 집계 규칙

- **텍스트 필드**: attack_type·geo.country·client_ip·url·user_agent는 .keyword를 붙여야 Terms가 동작한다.
- **숫자 필드**: http.status·http.bytes는 그대로 집계·범위 지정에 쓴다.
- **날짜 필드**: @timestamp는 그대로 Date Histogram의 축으로 쓴다.

S4
시각화

Date Histogram — 시간 버킷

> @timestamp를 시간 간격으로 묶으면 추이와 스파이크가 보인다.

시간 버킷이란

- **구간 분할**: @timestamp를 분·시간 같은 일정 간격으로 나눠 각 구간의 건수를 센다.
- **interval**: 자동으로 두면 화면 범위에 맞게 조정되고, 수동으로 1m·1h처럼 고정할 수도 있다.
- **축 배치**: 보통 X축에 두어 시간 흐름을, Y축 Count로 양을 표현한다.

활용

- **트래픽 추이**: 평상시 흐름을 파악해 야간·특정 시각의 이상 증가를 판단한다.
- **공격 스파이크**: event:attack 필터와 함께 쓰면 공격이 몰린 시간대를 즉시 찾는다.

S4
시각화

시각화 컴포넌트 설정표

랩에서 만드는 대표 컴포넌트의 설정과 보는 것을 한 표로 정리한다.

컴포넌트	Metric / Bucket 설정	보는 것
Metric 공격 건수	필터 event:attack · Metric = Count	총 공격 수 한 눈에
Pie / Donut 상태코드	Split slices → Terms → http.status	2xx/4xx/5xx 비율
Vertical Bar 공격 유형	필터 event:attack · Terms → attack_type.keyword	가장 많은 공격 벡터
Line 트래픽 추이	X-axis → Date Histogram → @timestamp	시간대별 스파이크
Data Table 상위 IP	필터 event:attack · Terms → client_ip.keyword (Size 10)	집중 공격 출발지
Heat Map	X=geo.country.keyword, Y=http.status, 값=Count	국가×상태코드 밀도
Tag Cloud	Terms → user_agent.keyword	많이 쓰인 UA(도구)

S4
시각화

Metric과 Pie

> 단일 KPI는 Metric, 구성비는 Pie로 표현한다.

항목	설정
Metric 용도	단일 KPI 카드
Metric 설정	Metric = Count
Metric 예	총 공격 수
Pie 용도	구성 비율
Pie 설정	Split slices = Terms
Pie 예	2xx/4xx/5xx 비율

읽는 법

- **Metric**: 한 개의 큰 숫자로 핵심 지표를 카드처럼 보여준다.
- **Pie**: 상태코드처럼 전체 대비 각 값의 비중을 조각으로 표현한다.
- **필터**: event:attack 필터로 범위를 공격 이벤트로 한정한다.

S4
시각화

Bar와 Line

> 순위는 Bar, 시간 추이는 Line으로 표현한다.

항목	설정
Vertical Bar	X = Terms attack_type.keyword
Bar 용도	값 순위 비교
Horizontal Bar	라벨이 긴 경우
Line	X = Date Histogram
Line 용도	시간대별 추이
다중 시리즈	Split series로 분리

읽는 법

- **Bar**: 막대 높이로 상위 공격 벡터의 순위를 비교한다.
- **Line**: 선의 오르내림으로 시간대별 스파이크를 확인한다.
- **Split series**: event별로 선·막대를 나눠 겹쳐 본다.

S4
시각화

Data Table와 Heat Map

> 목록은 Data Table, 두 축 교차 밀도는 Heat Map으로 본다.

Data Table

- **표 형태**: Terms 버킷 결과를 값과 Count의 표로 보여주어 정확한 수치를 읽는다.
- **상위 공격자**: client_ip.keyword를 Size 10으로 두면 집중 공격 IP 목록이 만들어진다.
- **정렬**: Count 내림차순으로 두어 가장 활발한 출발지를 위로 모은다.

Heat Map

- **두 축 교차**: X=geo.country.keyword, Y=http.status로 두 축을 교차시킨다.
- **밀도 표현**: 값=Count로 두면 국가×상태코드 조합의 밀도가 색 진하기로 드러난다.

S4
시각화

Region Map — 국가별 접속 지도

> geo.country 값을 세계 국가 경계에 조인해 접속량을 색으로 본다.

```
① Visualize → Create → Region Map
② 인덱스:  logs-demo-*

③ Buckets → Shape field
   Aggregation:  Terms
   Field:  geo.country.keyword
   Size:  50

④ Options → Layer Settings
   Vector map:  World Countries
   Join field:  ISO 3166-1 alpha-2 (iso2)

⑤ Color schema:  Yellow to Red
⑥ 상단 Update → Save
```

조인의 핵심

- **iso2 조인**: geo.country 값이 KR·US·CN 같은 ISO2 코드라 iso2 조인 필드로 국가 경계와 맞춘다.
- **색 표현**: 접속량이 많을수록 진한 색으로 칠해져 어느 국가가 몰리는지 보인다.
- **이상 탐지**: 서비스 대상국이 아닌 지역에 색이 집중되면 지오펜싱을 검토한다.

S4
시각화

Tag Cloud와 기타 컴포넌트

> 목적에 맞는 보조 컴포넌트를 골라 관제 화면을 풍부하게 한다.

보조 컴포넌트

- **Tag Cloud**: user_agent.keyword 빈도를 워드클라우드로 그려 많이 쓰인 도구·UA를 부각한다.
- **Gauge / Goal**: 임계치 대비 현재 값을 게이지로 표현해 목표 달성 여부를 본다.
- **Coordinate Map / Maps**: 좌표 기반 지도로 위경도 단위 접속 위치를 표시한다.

선택 기준

- **무엇을 보는가**: 추이는 Line, 비율은 Pie, 순위는 Bar·Table, 밀도는 Heat Map으로 정한다.
- **심화 시계열**: TSVB·Timelion은 다중 지표·수식이 필요한 시계열 분석에 쓴다.

S4
시각화

저장·공유와 예시 활용

> 공유 환경 규칙을 지키고 제공된 예시로 설정을 빠르게 익힌다.

저장 규칙

- **이니셜 접두**: 공유 환경이므로 이름 앞에 본인 이니셜을 붙인다(예 kim-공격유형).
- **재사용**: 저장한 시각화는 다음 세션의 대시보드에 그대로 얹어 재사용한다.
- **구분**: 이름만으로 누가 만든 무엇인지 알 수 있게 내용을 명확히 붙인다.

제공 예시 활용

- **[랩] 예시**: 공격 건수·상태코드·공격 유형·국가별 접속·트래픽·상위 IP 예시가 이미 저장돼 있다.
- **설정 참고**: 목록에서 [랩] 항목을 열어 Metric·Bucket 설정을 그대로 참고해 따라 만든다.

S4
시각화

세션 정리와 다음 예고

> 컴포넌트와 집계를 하나로 묶어 시각화의 핵심을 정리한다.

기억할 것

- **컴포넌트 선택**: 추이는 Line, 비율은 Pie, 순위는 Bar·Table, 밀도는 Heat Map, 분포는 Region Map으로 고른다.
- **Terms와 .keyword**: 텍스트 필드를 나눌 땐 반드시 필드.keyword를 써야 Terms 집계가 동작한다.
- **Date Histogram**: @timestamp를 시간 버킷으로 묶어 추이와 공격 스파이크를 읽는다.

다음 예고 — 세션 5

- **대시보드 조립**: 저장한 시각화들을 한 화면에 모아 SOC 보안 관제 대시보드를 만든다.
- **실시간 관제**: 시간 범위와 Auto refresh, 공통 KQL 필터로 모든 패널을 동시에 제어한다.